Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-018-1065-z

ORIGINAL RESEARCH

Affective recognition from EEG signals: an integrated data-mining approach

Fabio Mendoza-Palechor[1] · Maria Luiza Menezes[2] · Anita Sant'Anna[2] · Miguel Ortiz-Barrios[3] · Anas Samara[4] · Leo Galway[4]

Received: 26 April 2018 / Accepted: 22 September 2018

Abstract
Emotions play an important role in human communication, interaction, and decision making processes. Therefore, considerable efforts have been made towards the automatic identification of human emotions, in particular electroencephalogram (EEG) signals and Data Mining (DM) techniques have been then used to create models recognizing the affective states of users. However, most previous works have used clinical grade EEG systems with at least 32 electrodes. These systems are expensive and cumbersome, and therefore unsuitable for usage during normal daily activities. Smaller EEG headsets such as the Emotiv are now available and can be used during daily activities. This paper investigates the accuracy and applicability of previous affective recognition methods on data collected with an Emotiv headset while participants used a personal computer to fulfill several tasks. Several features were extracted from four channels only (AF3, AF4, F3 and F4 in accordance with the 10–20 system). Both Support Vector Machine and Naïve Bayes were used for emotion classification. Results demonstrate that such methods can be used to accurately detect emotions using a small EEG headset during a normal daily activity.

Keywords Affective recognition · Statistical features · Affective computing · Electroencephalogram (EEG) · Data Mining (DM)

✉ Miguel Ortiz-Barrios
mortiz1@cuc.edu.co

Fabio Mendoza-Palechor
fmendoza1@cuc.edu.co

Maria Luiza Menezes
maria.menezes@hh.se

Anita Sant'Anna
anita.santanna@hh.se

Anas Samara
samara-A@ulster.ac.uk

Leo Galway
l.galway@ulster.ac.uk

[1] Department of Electronic and Systems Engineering, Universidad de la Costa CUC, Barranquilla, Colombia

[2] Center for Applied Intelligent Systems Research, Halmstad University, Halmstad, Sweden

[3] Department of Industrial Management, Agroindustry and Operations, Universidad de la Costa CUC, Barranquilla, Colombia

[4] School of Computing, Computer Science Research Institute, Ulster University, Belfast BT37 0QB, UK

1 Introduction

Emotions are defined as a set of stimuli that any person feels when facing different past or present events. In this regard, emotions are also considered as the body's responses to such stimuli: physiological excitement, expressive conduct and conscious experience as stated by Barrett et al. (2016). Emotions play an important role in human interactions and decision making. Therefore, the ability to automatically detect emotions is important for any artificial system that interacts with humans. Consequently, in order to progress towards a more purposeful a beneficial form of human–machine interaction.

Data mining (DM) and machine learning techniques can be used to create models for automatic affective recognition. DM-based affective recognition may be useful for identifying specific behaviors and attitudes evidenced by people, identifying lifestyles and supporting decision-making in both medicine and education fields. Several authors like Koelstra et al. (2012), Soleymani et al. (2012), Liu and Sourina (2013), Wu et al. (2016), Chatchinarat et al. (2017), Katsigiannis and Ramzan

Journal : Large 12652	Article No : 1065	Pages : 20	MS Code : AIHC-D-18-00347	Dispatch : 26-9-2018

(2017), Shu and Wang (2017), Zhong and Jianhua (2017) and Menezes et al. (2017) have proposed DM techniques for affective recognition.

DM methods are dependent on good quality datasets for training models. In order to contribute to the development of good affective recognition algorithms, benchmark datasets have been created and are maintained by different research teams (Parsons and Rizzo 2008; Koelstra et al. 2012; Soleymani et al. 2012; Liu and Sourina 2013; Katsigiannis and Ramzan 2017). A summary of these datasets is presented in Table 1. Most often, benchmark datasets make use of 32 electrodes placed in accordance with the 10–20 system (Abadi et al. 2015). In some cases, more portable devices like the *Emotiv* (Wu et al. 2016) are used. In order to evoke emotional stimuli, participants are often shown videos or images and then asked to rate their emotional response in terms of valence and arousal with the help of a self-assessment maniquin. Such studies aim at acquiring high quality data with reliable ground truth; but are not representative of normal daily activities. Considering the aforementioned facts, our study aims to evaluate if current affective recognition models and strategies can be applied to data collected in less controlled experiments that simulate activities typical of daily living, in particular, using a personal computer to complete several common computer-based tasks. Feature extration techniques as well as machine learning models are used to create an affective recognition model. Model performace is evaluated based on self-reported ground truth.

The remainder of this paper is organized as follows: Sect. 2 introduces previous related work on emotion recognition. Methods used in the present work are explained in Sect. 3. In Sect. 4, the results are shown and analyzed. Finally, Sect. 5 presents conclusions.

2 Background

2.1 Benchmark datasets

Currently, various input modalities exist that can be utilized to acquire information about users and their emotions. More commonly, audiovisual communication, such as eye gaze tracking, facial expressions, body movement detection, and speech and auditory analysis may be employed as input modalities. Furthermore, physiological measurements using sensor signals, such as EEG, galvanic skin response, and electrocardiogram can also be utilized. However, the use of EEG as an input modality has a number of advantages that make it potentially suitable for use in real-life tasks including its non-invasive nature and relative tolerance to movement. EEG can be used as a standalone modality as well as combined with other biometric sensors. Considering the reported literature, many efforts have been made by different authors to contribute to the affective recognition field and multiple datasets have been built to be effectively used when creating new classifiers.

The creation of accurate machine learning models from EEG data depends on the quality of the data that is used. In order to further develop in this field, several researchers have created benchmark databases. Koelstra et al. (2012) proposed a dataset called "DEAP", which consists of EEG signals and peripheral physiological signals derived from 32 participants. These signals were recorded while the applicants viewed 40 1-min musical videoclips. In this work, a high positive correlation was found between liking/dominance and valence since people like music that gives empowerment sensations. On the other hand, a moderate positive correlation was detected between liking/dominance and arousal (Koelstra et al. 2012).

A multimodal database called MANHOB-HCI, which is used for recognizing human affect and implicit labeling,

Table 1 Datasets for affective recognition

Database	Description
DEAP (Koelstra et al. 2012)	32 participants with each one seeing 40 one-minute videos and the use of electrodes in different brain regions for data collection
MAHNOB-HCI (Soleymani et al. 2012)	27 people with each one initially seeing 20 videos. Then, a data-collection process took place with the participants observing brief video clips and images and using electrodes in different brain regions
Liu and Sourina (2013)	14 participants whose data were stored and used for affective recognition. In this experiment, audio and visual stimulus were implemented and the data-collection process was conducted with the support of the *Emotiv* device
DECAF (Parsons and Rizzo, 2008)	30 participants with each one seeing 40 one-minute musical video segments and 36 movie clips which allows to compare EEG and MEG modalities as well as analyzing the people stimulus when listening to music. This is also used for affective recognition
DREAMER (Katsigiannis and Ramzan 2017)	23 participants and the integration of EEG and ECG signals

Journal : **Large 12652**	Article No : **1065**	Pages : **20**	MS Code : **AIHC-D-18-00347**	Dispatch : **26-9-2018**

was built by Soleymani et al. (2012). To do this, it was necessary to record the responses to emotion stimuli aiming at identifying the emotions of 27 participants. The dataset gathers information on face poses, audio signals, eye gaze and peripheral physiological signals. The experimentation was comprised of two phases. First, the participants saw 20 videos in order to detect their emotions through the use of excitement, valence, dominance, predictive ability and emotional keywords. In the second phase, the participants visualized short videos and images which were presented once with and without correct labeling. This was assessed in order to evidence their agreement or disagreement with the respective labeling. Here, the authors used a hidden Markov model for classifying the sequence of facial expressions in accordance with the correction of the previously shown labels. Furthermore, the classification process was evaluated by applying cross-validation methods (Soleymani et al. 2012).

The use of EEG for affective recognition was also expressed by Liu and Sourina (2013) as, by using electroencephalogram (EEG), is an aspect of interest for the research community. Therefore, the above-mentioned authors created a dataset for emotion classification using audio and visual stimulation during the experimentation process. The stimulus is selected from the International Affective Digitized Sound Systsme (IADS) and the International Affective Picture System (IAPS) datasets. For dataset construction, the *Emotiv* device was employed, to collect the response of 14 participants. The stimuli are classified by the participants considering the arousal, valence and dominance levels. In addition, the authors analyze the correlation degree between different EEG frequency bands and affect assessment. The approach proposed by the authors consists of two phases. Initially, there is an extraction process using a sliding window followed by a data classification algorithm applying Support Vector Machine (SVM). Finally, the presented method is able to recognize eight emotions: joy, surprise, satisfaction, protected, angry, frightened, unconcerned and sad. The best accuracy result for classification of 8 emotions is 53.7% by using four electrodes whilst, 87.02% is the best outcome when recognizing two emotions under the same number of electrodes (Liu and Sourina 2013).

DECAF, a multimodal database that allows researchers to de-codify the physiological user responses to multimedia content was presented by Abadi et al. (2015). Correspondingly, the DECAF dataset contains brain signals that are obtained by using a Magnetoencephalogram (MEG) sensor that requires low physical contact with the user's scalp. Moreover, DECAF (Parsons and Rizzo 2008) contains emotional implicit and explicit reactions from 30 participants seeing 40 segments of one-minute musical videoclips. This facilitates the comparisons between EEG and MEG modalities. In addition to the MEG, the DECAF dataset, contains synchronized Near Infrared Reflectance (NIR) face videos, Horizontal Electro-Oculogram (HEOG), Electro-oculogram (OCG), Electrocardiogram (ECG) and peripheral physiological responses of trapezoid electromyogram (TEMG).

Another multimodal database, DREAMER, comprising information on EEG and ECG signals from 23 people was provided by Katsigiannis *et al*. (2017). The stored information corresponds to audiovisual stimulus where the affective state was analyzed and compared to valence, arousal and dominance. Every signal was collected by using portable devices and wearable sensors that allow the use of affective computing methods in day-to-day applications. The authors propose the use of Support Vector Machine (SVM) for affective recognition based on EGG and ECG (Katsigiannis and Ramzan 2017). Table 1 summarized the available databases that were created for affective recognition.

2.2 Related work

As stated by Chatchinarat et al. (2017), the affective recognition and classification based on EEG signals are widely studied because of their potential benefits for both healthcare and entertainment fields. In this regard, different methods can be used for the classification process; for instance, SMV may be combined with a decision tree approach to achieve better accuracy results compared to those reported in the literature.

In performing affective recognition from EEG signals, it is not common to consider multiple subjects and individual patterns for each subject simultaneously, as expressed by Wu et al. (2016). They presented a novel approach for affective recognition where subjects, or a set of them, are used as contributors of relevant information. In their work, five frequency attributes were extracted from each EEG signal. These parameters were selected by carrying out statistical tests. Finally, the proposed method evidenced that two three-node Bayesian networks can be used to capture probability distribution functions for emotion labeling.

By contrast, Shu and Wang (2017) established that the dependence among multiple physiological signals is the cornerstone of multimodal affective recognition; however, it has not been exploited entirely. Consequently, this study proposed to use the Restricted Boltzmann Machine (RBM) for dependency modeling. Specifically, the RBM visible nodes represent the EEG and the peripheral physiological signals; hence, the links between visible and hidden nodes identify the intrinsic interlinkages among multiple signals. The authors applied SVM for affective recognition from the generated attributes.

Combining machine learning and DM approaches is considered by Zhong and Jianhua (2017) to be an interesting proposal for research due to the use of physiological data such as EEG signals for affective recognition based on physiological data. Particularly, the classification models

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018

can be learned from heterogeneous attributes. The set of subject-independent EEG features using *transfer recursive feature elmination* (T-RFE), which allows obtaining the subset of optimal characteristics. The authors used DEAP as a data source in conjunction with the *linear square support vector machine* (LSSVM) as a base for selecting the EEG attributes.

Menezes et al. (2017) used the DEAP dataset for emotion classification from several features. Reasonable classification accuracies for Valence and Arousal were obtained via calculating feature vectors based on statistical measurements, band power from α, β, δ, and θ waves of the EEG signal.

Considering the reported literature, statistical methods have been widely used to design and develop smart tools for affective recognition as well as the identification and extraction of attributes. A statistical feature is a distinctive characteristic of a dataset obtained from different types of mathematical transformation (Barrios and Jiménez 2015). Particularly, it is used for supporting human emotion classification due to the notorious difficulties identified when using bio-signals. The research findings suggest that, once the signals are pre-processed, brainwaves can be successfully characterized using statistical features (Jerritta et al. 2011). This is useful when considering that a feature must demonstrate high stability in order to be accepted for clinical use (Lan et al. 2016). Algorithms based on statistical features have become the most used feature extraction techniques (Schaaff and Schultz 2009; Chai et al. 2010; Mampusti et al. 2011; Bastos-Filho et al. 2012) and several authors have attempted to find the attributes providing the highest affective recognition accuracy. Subasi (2007) used four statistical features to represent the time–frequency distribution of the EEG signals (diagnosis of epilepsy): Mean of absolute values of the coefficients in each sub-band (1), average power of the wavelet coefficients in each sub-band (2), standard deviation of the coefficients in each sub-band (3) and ratio of the absolute mean values of adjacent sub-bands (4). Features (1) and (2) were then combined to denote the frequency distribution of the signal whilst (3) and (4) were employed to estimate the number of changes in the frequency distribution.

Murugappan et al. (2008a, b) proposed an affective recognition system from EEG signals and computed three statistical features for classifying human emotions: energy, recoursing energy efficiency (REE) and root mean squares (RMS). Specifically, REE has efficiently clustered the emotions by achieving the performance goal (Murugappan et al. 2010). Meanwhile, Chai et al. (2010) proposed a statistics-based system for human emotion classification by using EEG. In this study, six statistical features were computed: means of the raw signals (1), standard deviation of the raw signals (2), means of the absolute values of the first differences of the raw signals (3) means of the absolute values of the first differences of the normalized signals (4) means of the absolute values of the second differences of the raw signals (5) and the means of the absolute values of the second differences of the normalized signals (6). These statistics have been also used in Picard et al. (2001), Maaoui and Pruski (2010), Lan et al. (2014), Menezes et al. (2017) and Nugent et al. (2016). Particularly, Lan et al. (2014) found that the standard deviation and the mean of the absolute values of the second differences of the normalized EEG proved to be satisfactory regarding intra-class correlation coefficient (ICC). Furthermore, a combination of these measures, was employed. In this respect, the vector (3)–(5) produced the highest rate of correct classification (95%) and 12.68 s were consumed for training. However, 100% correct classification was only achieved for the emotion "sadness". In this sense, all the testing inputs for "sadness" were correctly identified as "sadness" Consequently, more work should be emphasized in augmenting the effectiveness of algorithms in recognizing a higher number of emotions as well as reducing the processing time required by the algorithm in producing positive results. Another example can be found in Murugappan et al. (2009) who investigated the possibility of using visual and audiovisual stimuli for detecting human emotion by measuring EEG. Herein, two statistical features were extracted for each channel on alpha frequency band: energy and power.

Statistical features comprising the selected mean, median, standard deviation, skewness and kurtosis were employed by Islam et al. (2013) to represent the largest dispersion in different mental states and to help assess different human emotions. In this study, the skewness of EEG signals determined the peakedness in the state of relaxing, thought, memory, motor action, fear, pleasant state and enjoying music. In addition, it provided further information of the brain or cognitive functions in different frequency components.

When combined with other methods, statistical features can also provide very good results as stated by Rizon et al. (2008) who used four statistical measures (energy, normalized energy, entropy and power) combined with "db4" wavelet function. The results demonstrated that this technique performed well in classifying the emotions on an optimal set of channels proposed by the asymmetric ratio-based channel selection method. Also, Liu and Sourina (2014) integrated statistical parameters with Fractal dimension features to improve accuracy and generate adequate computational time. The results evidenced that two emotions can be recognized with the best average accuracy of 87.02% when using 4 four electrodes.

Wang et al. (2011) concluded that the classification performance using all statistical features is evidently better than those based on individual features under the same conditions. In this regard, Kim and André (2008) investigated the potential of physiological signals as reliable channels for

Journal : Large 12652	Article No : 1065	Pages : 20	MS Code : AIHC-D-18-00347	Dispatch : 26-9-2018

affective recognition. Herein, the authors used extended Linear Discriminant Analysis (plDA) to classify four musical emotions (positive/high arousal, negative/high arousal, negative/low arousal, and positive/low arousal). An improved recognition accuracy of 95% and 70% for subject-dependent and subject-independent classification, respectively, were achieved. Likewise, Vijayan et al. (2015) proposed a novel approach based on statistically weighed autoregressive modeling of EEG for the classification of human emotions. The algorithm was evidenced to be superior to other related techniques since it provided a classification accuracy of 94.097%. Also, it is useful to make the emotion classification process simpler. In this respect, Wang and Sourina (2013) applied Principal Component Analysis (PCA) combined with the six measures proposed by Picard et al. (2001) in order to eliminate redundant information within the extracted statistical features, which may result in a reduction with respect to the initial number of features. Similarly, Atkinson and Campos (2016) used the minimum–Redundancy–Maximum–Relevance (mRMR) method (Wu et al. 2010; Liu et al. 2010) to select a relevant set of parameters so that further classification can be more accurate. It was demonstrated that mRMR outperformed other state-of-the-art techniques.

As concluded by Jerritta et al. (2011), real-time affective recognition using physiological signals is still in its early stages of growth. As emotions are highly subjective, an overall framework for classifying all the basic emotions remains a challenge. Despite the studies conducted for this purpose, it is still necessary to develop efficient feature extraction algorithms using a different set of statistical parameters for improving the emotion classification rate. In addition, it was established that classification based on arousal and valence values proved to be rather interesting. Another finding is that there is no comparative study determining the statistical correlation between different affective states and the waves derived from EEG signals.

In light of these, the conducted literature review showed that the studies concentrated on the use of kurtosis, skewness and median are largely limited. Therefore, we implemented these parameters in this study in conjunction with other traditional measures (i.e. mean and standard deviation) in order to explore their effectiveness when classifying emotions and to subsequently provide features that can be used in realistic daily living scenarios.

3 Methods

3.1 Dataset preparation and analysis

The data-collection process included the following sensing modalities: (1) depth camera (Intel Real-Sense 3D), (2) eye tracker (eye tribe tracker), (3) Emotiv EPOC headset to record EEG behavior during the task attempts, use (4) microphone to record participant voice while he/she implemented the Talk Aloud Protocol (TAP). In this study, however, we focus on the analysis of the EEG signal only. The data collection study was undertaken at the Artificial Intelligence Application Research Group (AIARG) lab at Ulster University, Belfast, UK. The resulting number of instances per participant $n_p \sim N[6680;5056]$ and the size of the final dataset was 140724 (including 132 features). The study was approved by the Ulster University Ethics Filter Committee (FCE 20160419 16.24). During the study participants were asked to perform four computer-based tasks using common computer software while seated at a desktop-based personal computer.

The set of four tasks with associated sub-tasks was as follows:

1. Basic operating system task (adjust desktop computer system):
   a. Change Desktop background, desktop resolution, screen saver and, create/move/delete folders
   b. Change regional settings, time zone, currency and add new language
2. Online shopping task find tablet PC online using preferred browser:
   a. With a screen size equal to or greater than 7 inches and where the price is less than £50
   b. In addition to (a), where the tablet has 16 GB storage and a camera equal to or greater than 5MP
3. Excel spread sheet tasks (manipulate the pre-populated spreadsheet):
   a. Insert a new record into the spreadsheet, sort the names into ascending order and verify that the actions were applied
   b. Calculate the average and create a line chart from the data
4. Game-based tasks: participants were asked to play Pacman (Deluxe Pacman 2) with two levels of difficulty:

For each task, a maximum time limit of two minutes was given, with the exception of the game-based task, which was limited to three minutes with an initial period of familiarization prior to starting the task. Tasks were presented in a random sequence in order to eradicate bias.

Initially, each participant was given an information sheet describing the flow of the study, along with the equipment to be used. Following this, consent for participation was given (if agreed), and both Emotiv EPOC and Eye Tribe Tracker setup and calibrated. The participant the commenced the first of the four selected tasks according to instructions given in an accompanying task sheet. Upon completion of a task, the task time and completion state were determined, and the participant was asked to self-report on his/her feelings regarding the task using the Self-Assessment Manikin (Bradley and Lang 1994) shown in Fig. 1, in addition to annotating selected facial images acquired during the task. A minimum of three facial images captured during a task were chosen by the participant, whereby Valence and Arousal values from the range [1–9] were utilized in concert with the Self-Assessment Manikin to annotate the selected facial images. This self-reporting process was repeated after each of the four tasks. The information on perceived Valence and Arousal by each participant for each task will subsequently be used for further analysis.

All participants were either staff or students at Ulster University, however, due to the overarching focus of the study no demographic information was recorded. In addition, there was no pre-defined exclusion criteria, hence participant's prior computer experience could vary from novice to expert. The subsequent dataset obtained includes information on the emotional states of 22 participants, leading to a self-reported Valence and Arousal values from the Self-Assessment Manikin post-task, and a total of 304 instances (on average) of perceived Valence and Arousal from the selected facial images acquired during each task.

3.2 Support vector machine (SVM)

Commonly used to solve prediction and classification problems in an efficient way due to its automatic learning system. They are based in the statistic learning system developed by (Niedermeyer and Da Silva 1993), when a mathematic model is proposed for regression and classification problems (Parsons and Rizzo 2008).

Other authors mention that SVM is a margin classifier that gets trained by a dataset with feature vectors. SVM tries to find an optimal limit that separates two classes with different feature vectors with a maximal margin (distance between optimum hyperplane and the nearest vector). To make classification of an inseparable dataset, a nonlinear SVM projects a feature vector in a high dimensional space using a kernel function such as radial basis kernel function (Botella et al. 2004).

The construction of SVM is based on transforming or projecting a dataset in a given n dimension to higher dimension space applying a kernel function—kernel trick. From this new space created, the data is operated as a linear problem, solving it without considering the data dimensionality (Brahnam and Jain 2010). AQ3

Some advantages of SVM are: First, it has a solid mathematics foundation. Second, it has the concept of structural risk minimization (Hodges et al. 2001; Glantz et al. 2003), that translates into the minimization of the probability of a wrong classification on new examples. This case is very common when there are too few data for training. The third advantage relies on the availability of powerful tools and algorithms to find the solution in fast and efficiently (De la Hoz et al. 2014; Bekele et al. 2016).

3.3 Naïve Bayes

Bayesian networks are considered an alternative to classic expert systems oriented to decision making and prediction under uncertainty in probabilistic terms (Picard et al. 2004). In Bransford et al. (1999) and Ip et al. (2011), a structure composed of four levels is used. At the highest level would be a set of variables mapped by nodes and arrows that relate with influence terms. In the next level, you would find the levels or states, also known as *state space* that can take each of the model variables (Ontiveros-Hernández et al. 2013). In third place, you can find a set of conditional probability functions, one for each node, and represents the probability of occurrence of each state of the variable conditioned to possible values. At the lowest level, is a set of algorithms

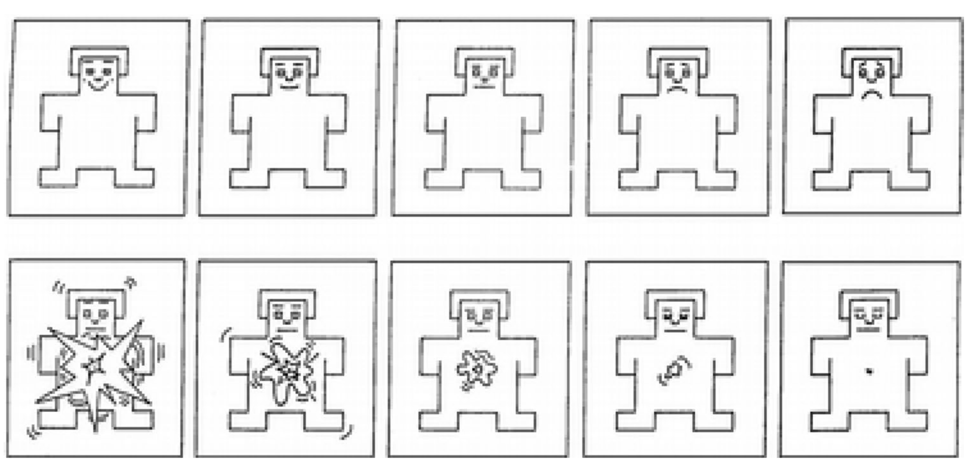


Fig. 1 Self-Assessment Manikin (SAM), used by participants to assess level of Valence and Arousal

Journal : **Large 12652**	Article No : **1065**	Pages : **20**	MS Code : **AIHC-D-18-00347**	Dispatch : **26-9-2018**

that would allow the network to recalculate the probabilities assigned to each of the levels when some evidence from the model is known.

4 Description of proposed methodology

4.1 Selection of EEG channels

Research has evidenced that the frontal lobe is key when measuring emotions. It has significant activity during the experience of emotions, affective reactions and emotion regulation (Konstantinidis et al. 2012). As a first experiment and in order to continue the work in Menezes et al. (2017), we chose to use only the EEG signal from positions Af3, Af4, F3 and F4 (related to prefrontal cortex and frontal lobes), as seen in Fig. 2. These signals were acquired with an Emotiv EPOC headset. This selection also aims to study the effectiveness of a reduced number of electrodes to analyze affective states. This would provide a simpler and more user-friendly data acquisition for future use on the wild and in real-time situations.

4.2 Bandwave extraction

Parks–McClellan algorithm and Chebyshev Finite Impulse Response filter were applied to the EEG signal in order to obtain the brainwaves Delta (δ), Theta (θ), Alpha (α) and Beta (β). The frequency ranges to obtain each wave were as follows: Delta (δ) from 0.5 to 4 Hz; Theta (θ) from 4 to 8 Hz; Alpha (α) from 8 to 12 Hz; and Beta (β) 12 to 30 Hz (Menezes et al. 2017).

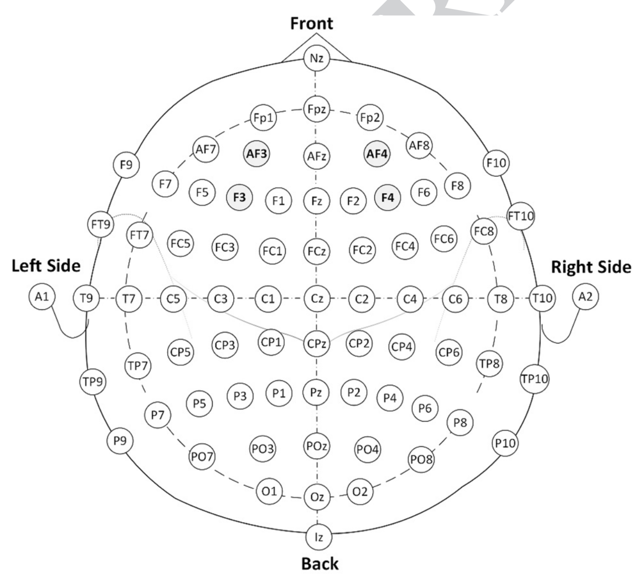


Fig. 2 Af3, Af4, F3 and F4 positions selected according to the 10–20 system

4.3 Feature extraction

During the cleaning process, the signals were downsampled to 125 Hz and high-pass filtered with a cut-off frequency of 2 Hz by using Matlab. Different kinds of features were then calculated from EEG signals. Here, statistical and power-band parameters were considered. Such measures and the construction of feature vectors are further explained below. In this case, there is not any data mixing the four electrodes during the extraction of the characteristics.

4.3.1 Statistical features

Seven statistical parameters were calculated for each of the signals as follows. Let the data from the EEG headset be represented by X. This data includes four signals, one from each channel position ($AF3$, $AF4$, $F3$, $F4$ according to the 10–20 system). The signal from each channel was decomposed into four frequency bands: α, β, δ, and θ as explained above. For each participant, each observation corresponds to a task performed by the participant, so the data were segmented according to the duration of each of the tasks. X_{cp} is defined as the nth ($n = 1, \ldots, N$). sample (in time) for task c obtained from the p channel position. Here, N represents the length of the task. Statistical features were computed over a window (± 2 s) encompassing the entire task. In addition, $\mu_{x_{cp}}$ (refer to Eq. 1) and $\sigma_{x_{cp}}$ (refer to Eq. 2) are the mean and standard deviation of X_{cp} respectively, whilst the absolute average and deviation are $\left|\mu_{x_{cp}}\right|$ (refer to Eq. 3) and $\left|\sigma_{x_{cp}}\right|$ (refer to Eq. 4) correspondingly.

$$\mu_{x_{cp}} = \frac{1}{N}\sum_{n=1}^{N} X_{cp(n)} \tag{1}$$

$$\sigma_{x_{cp}} = \left(\frac{1}{N-1}\sum_{n=1}^{N}\left(X_{cp(n)}(-\mu_x\right)^2\right)^{1/2} \tag{2}$$

$$\left|\mu_{x_{cp}}\right| = \frac{1}{N}\sum_{n=1}^{N}\left|x_{cp(n)}\right| \tag{3}$$

$$\left|\sigma_{x_{cp}}\right| = \left(\frac{1}{N-1}\sum_{n=1}^{N}\left(\left|X_{cp(n)}\right| - \mu_{x_{cp(n)}}\right)^2\right)^{1/2} \tag{4}$$

In an effort to provide better accuracy measures, this study additionally focuses on the use of median (refer to

Eq. 5). Here, l is the lower class boundary of the median class; h denotes the size of the median class interval, f is the frequency of a median class and f_c represents the cumulative frequency preceding median class.

$$M_{x_{cp}} = l + \frac{h}{f}\left(\frac{N}{2} - c\right) \tag{5}$$

Other parameters of interest are skewness (refer to Eq. 6) and kurtosis (refer to Eq. 7). Particularly, the use of these features is largely limited in the reported literature. Therefore, we decided to explore their effectiveness in this study. In this regard, these measures may correlate with having an emotion and subsequently complement the traditional features (Eq. 1–4) proposed in other works.

$$SK_{x_{cp}} = \frac{\sum_{n=1}^{N}\left(X_{cp} - \mu_{x_{cp}}\right)^4}{(N-1)\sigma_{x_{cp}}{}^4} \tag{6}$$

$$k_{x_{cp}} = \frac{\sum_{n=1}^{N}\left(X_{cp} - \mu_{x_{cp}}\right)^3}{(N-1)\sigma_{x_{cp}}{}^3} \tag{7}$$

Although studies have expressed that there is a strong correlation between brainwaves and different affective states (Lin et al. 2010; Menezes et al. 2017), it is important to check that this is indeed true in our dataset. In this respect, the adjusted R^2 is calculated to estimate the percentage of response variable (both *Arousal* and *Valence*) variation that is explained by its relationship with the predictor variables but considering the number of predictors in the regression model. Furthermore, the predicted R^2 is computed to indicate how well the set of statistical features predict new responses of *Arousal* and *Valence*. Particularly, *adjusted* R^2 − *predicted* R^2 is of interest to determine whether the model is overfitted and adequate to provide valid predictions for new observations.

4.3.2 Affective state classification

The Circumplex Model of Affect is a valuable representation of all affective states. Herein, the emotions are classified along two independent dimensions (refer to Fig. 3): Arousal and Valence. Arousal, in the vertical axis, describes the extent to which an affect is correlated to an individual sensation of energy; whilst Valence, in the horizontal axis, represents the degree to which an emotion reveals a positive or negative state of mind (Gerber et al. 2008).

As the primary aim of this research is to correctly identify the human emotional states, the Circumplex Model of Affect was utilized. This is consistent with the recent findings from the neuroscience, behavioral and cognitive research fields (Pool et al. 2016; Binder et al. 2016; Desmet 2018). In this regard, the first step involved collecting the Arousal and Valence values (SAM scale) reported by the participants (Barakat and Bradley 2010). These values were later discretized using the tripartition and bipartition labeling schemes as follows: (1) Tripartition: Low [1.0–3.0], Medium [4.0–6.0] and High [7.0–9.0] whilst (2) Bipartition: Low [1.0–3.0] and High [7.0–9.0]. Finally, the EEG biosignals were classified through SVM (Liu and Sourina 2013; Chatchinarat et al. 2017; Menezes et al. 2017; Katsigiannis and Ramzan 2017) and Naïve Bayes (Kim et al. 2010; Jirayucharoensak et al. 2014). Naïve Bayes was selected due to: (1) its high computational efficiency, (2) versatility, (3) easiness of implementation, (4) high scalability, (5) low need of training data, (6) suitability for binary and multiclass classification problems and (7) capability of handling continuous and discrete data. On the other hand, SVM was chosen since: (1) it can avoid overfitting, (2) it is flexible due to the introduction of kernel, (3) it is robust against different outliers and model violations and (4) it learns with a small number of predictors.

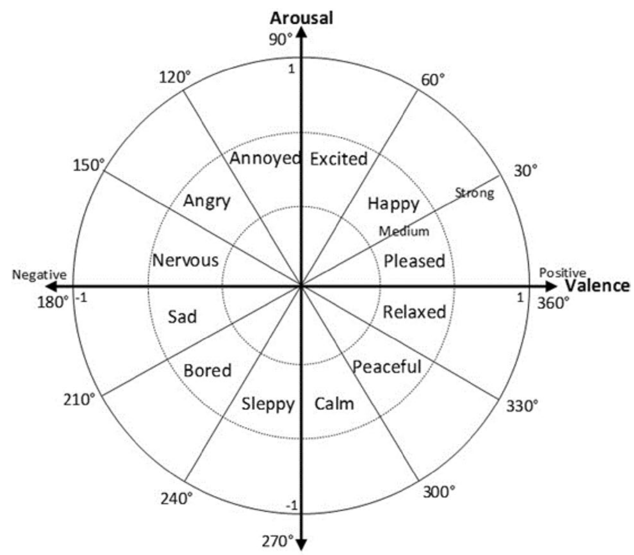


Fig. 3 The Circumplex Model of Affect (Gerber et al. 2008)

5 Results and discussion

The results of the statistical analysis conducted for feature extraction as well as the validation of data quality and attributes considered for affective recognition are presented in this section. In addition, the outputs of classification methods (SVM and Naïve Bayes) are also shown below.

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018

5.1 Statistical features

When performing the correlation analysis between the brainwaves (frequently categorized in four different frequency bands: α, β, δ, θ) and the affective states, the model evidenced a significant correlation between the response variables (*Arousal* and *Valence*) and the statistical features at a 5% significance level (p-value = 0). In this respect, the P-values of mean/absolute average (refer to Table 2), standard deviation/absolute deviation (refer to Table 3) and median/skewness/kurtosis (refer to Table 4) were estimated.

Specifically, it was found that $\mu_{x_{\beta(AF3)}}$ (0.041), $\mu_{x_{\theta(AF3)}}$ (0.003) and $k_{x_{\beta(f3)}}$ (0.029) were meaningfully related (P-value > 0.05) to *Arousal* values. This suggests that a quadratic model with the aforementioned statistical features may be appropriate (refer to Eq. 8) and there would therefore be more fit to train the models. The expression was established with the aid of Minitab 17 ® software by conducting a regression analysis.

$$Arousal = \left(0.613\mu_{x_{\beta(AF3)}} + 0.00309k_{x_{\beta(f3)}} - 0.4359\mu_{x_{\theta(AF3)}}\right)^2 \quad (8)$$

Likewise, it was concluded that $\sigma_{x_{\delta(AF4)}}$ (0.034) and $k_{x_{\delta(AF3)}}$ (0.048) are both significant to *Valence* values. After carrying out a regression study, a mathematical model with these parameters was achieved (refer to Eq. 9). Better fit and increased classification performance may be also expected when training the model.

Table 2 P-values for mean and absolute average of the brainwaves obtained from each position

Position (brainwave)	Arousal		Valence	
	$\mu_{x_{nc(p)}}$	$\mu_{x_{nc\ (p)}-ABS}$	$\mu_{x_{nc(p)}}$	$\mu_{x_{nc\ (p)}-ABS}$
AF3 (α)	0.675	0.691	0.932	0.450
AF3 (β)	0.041*	0.480	0.178	0.703
AF3 (δ)	0.062	–	0.600	–
AF3 (θ)	0.003*	0.433	0.258	0.075
AF4 (α)	0.913	0.480	0.449	0.466
AF4 (β)	0.672	0.125	0.621	0.130
AF4 (δ)	0.187	–	0.208	–
AF4 (θ)	0.174	0.570	0.066	0.723
f3 (α)	0.429	0.735	0.328	0.901
f3 (β)	0.790	0.633	0.800	0.620
f3 (δ)	0.081	–	0.584	–
f3 (θ)	0.986	0.855	0.573	0.311
f4 (α)	0.860	0.986	0.985	0.764
f4 (β)	0.872	0.254	0.888	0.080
f4 (δ)	0.076	–	0.545	–
f4 (θ)	0.422	0.999	0.143	0.541

Table 3 P-values for standard deviation and absolute deviation of the brainwaves obtained from each position

Position (brainwave)	Arousal		Valence	
	$\sigma_{x_{nc(p)}}$	$\sigma_{x_{nc\ (p)}-ABS}$	$\sigma_{x_{nc(p)}}$	$\sigma_{x_{nc\ (p)}-ABS}$
AF3 (α)	0.681	0.719	0.408	0.450
AF3 (β)	0.421	0.462	0.843	0.933
AF3 (δ)	0.071	–	0.174	–
AF3 (θ)	0.152	0.082	0.201	0.359
AF4 (α)	0.501	0.548	0.679	0.777
AF4 (β)	0.202	0.226	0.094	0.108
AF4 (δ)	0.336	–	0.034*	–
AF4 (θ)	0.524	0.521	0.589	0.327
f3 (α)	0.721	0.727	0.654	0.550
f3 (β)	0.662	0.651	0.575	0.626
f3 (δ)	0.252	–	0.562	–
f3 (θ)	0.730	0.752	0.372	0.392
f4 (α)	0.957	0.963	0.824	0.828
f4 (β)	0.255	0.282	0.060	0.074
f4 (δ)	0.133	–	0.399	–
f4 (θ)	0.933	0.846	0.935	0.714

$$Valence = \left(0.715k_{x_{\delta(AF3)}} + 0,000556\sigma_{x_{\delta(AF4)}} - 0.04318k^2_{x_{\delta(AF3)}}\right)^2 \quad (9)$$

It is of particular interest to note that kurtosis was found to be useful for both models. Therefore, it can be employed in future studies for supporting affective recognition activities. This should be complemented with the use of mean and standard deviation whose contribution is highly relevant upon correlating brainwaves and affective states. In contrast, median, skewness and absolute measures were not estimated as meaningful and were subsequently discarded in both Eq. 8 and Eq. 9. Another important finding is that most of the significant features are related to β (*Arousal*) and δ (*Valence*) frequency bands. Additionally, it was observed that *AF*3 was identified as the most contributing position for affective recognition.

Upon considering correlation measures, it can be appreciated that the model fits well ($R - sq(adj) = 94.90\%$) for *Arousal* and the predictive ability is highly satisfactory ($R - sq(pred) = 94.86\%$). Similarly, these metrics evidenced high correlation and prediction performance regarding *Valence* values with $R - sq(adj) = 85.08\%$ and $R - sq(pred) = 83, 10\%$. It is also important to consider that the difference between these parameters is non-significant: $0, 04\%$ *and* $1, 98\%$ for *Arousal* and *Valence* respectively. Hence, the models do not appear to be overfitted.

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018

5.2 Emotion classification

When recognizing different emotions, we used the *accuracy* and *recall* as key performance indexes for evaluating different classification methods. The true- and false-positive ratios were also considered for this purpose. In addition, stratified k-fold cross-validation was applied ten times (k = 10) in order to assess the classification performance. Specifically, the amount of processed data in bipartition approach was 82964; whilst, 140724 were used in tripartition labeling scheme. The number of data per subsample was then 8296.4 and 14072.4 for bipartition and tripartition correspondingly. This study aims to identify particular patterns regarding the features extracted from the EEG signals and their relation to different *Valence* and *Arousal* states. To do this, we implemented SVM and Naïve Bayes techniques. Furthermore, a bipartition and tripartition labeling scheme, as outlined in Sect. 4.3.2, was used for each of the affective domains.

Tables 4, 5, 6 and 7 present the results obtained from all the dataset instances, i.e., all the tasks performed by

Table 4 P-values for median, skewness and kurtosis of the brainwaves obtained from each position

Position (brainwave)	Arousal			Valence		
	$M_{x_{nc(p)}}$	$SK_{x_{nc(p)}}$	$k_{x_{nc(p)}}$	$M_{x_{nc(p)}}$	$SK_{x_{nc(p)}}$	$k_{x_{nc(p)}}$
AF3 (α)	0.414	0.227	0.572	0.674	0.876	0.561
AF3 (β)	0.749	0.431	0.029*	0.203	0.057	0.927
AF3 (δ)	0.365	0.114	0.753	0.995	0.856	0.048*
AF3 (θ)	0.236	0.523	0.103	0.872	0.208	0.639
AF4 (α)	0.427	0.933	0.885	0.849	0.967	0.403
AF4 (β)	0.657	0.385	0.481	0.314	0.967	0.150
AF4 (δ)	0.637	0.771	0.934	0.271	0.053	0.843
AF4 (θ)	0.839	0.439	0.752	0.768	0.630	0.391
f3 (α)	0.229	0.785	0.363	0.212	0.691	0.682
f3 (β)	0.570	0.347	0.799	0.133	0.130	0.380
f3 (δ)	0.175	0.283	0.102	0.149	0.230	0.593
f3 (θ)	0.244	0.992	0.259	0.170	0.832	0.114
f4 (α)	0.572	0.295	0.211	0.799	0.710	0.963
f4 (β)	0.506	0.196	0.671	0.224	0.290	0.368
f4 (δ)	0.082	0.169	0.459	0.627	0.326	0.817
f4 (θ)	0.726	0.784	0.700	0.492	0.739	0.244

Table 5 Results of classification process using tripartition labeling scheme (statistical and powerband parameters)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	78.0	70.6	70.6	4.4	76.8	67.8	67.8	6
	Medium	80.7	82.7	82.7	13.7	81.5	85.2	85.2	17
	High	78.7	80.0	80.0	15	74.9	76.2	76.2	11.2
Naïve Bayes	Low	22.7	86	86	65.4	25.9	86.9	86.9	73.1
	Medium	63.1	11.6	11.6	4.7	56.8	20	20	13.3
	High	50.9	29.1	29.1	19.4	67.4	16.1	16.1	3.4

Table 6 Results of classification process using bipartition labeling scheme (statistical and powerband parameters)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp Rate
Support vector machine	Low	92.4	92.1	92.1	3.4	91.8	78.5	78.5	5.2%
	High	96.5	96.6	96.6	7.9	85.5	94.8	94.8	21.5%
Naïve Bayes	Low	37.8	88.3	88.3	64.8	46.9	95	95	80.3%
	High	87.1	35.2	35.2	11.7	84	19.7	19.7	5%

Table 7 Results of classification process using tripartition labeling scheme (powerband parameters)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	0	0	0	0	37.5	0.5	0.5	0.3
	Medium	49.9	69.3	69.3	48.2	53.1	79.6	79.6	61.6
	High	57.9	61.1	61.1	30.7	41.2	40.1	40.1	25.1
Naïve Bayes	Low	31.7	8.9	8.9	4.3	25.9	86.9	86.9	73.1
	Medium	56.2	6.6	6.6	3.6	56.8	20	20	13.3
	High	41.7	91.9	91.9	88.8	67.4	16.1	16.1	3.4

participants, by using SVM and Naïve Bayes methods under a tripartition scheme. Particularly, Table 5 compares the two methods (SVM and Naïve Bayes) in terms of all the attributes (statistical and powerband parameters) relating to the extracted brainwaves (α, β, δ, θ). After conducting a paired sample t test from the results of Table 5, the p-values were found to be 0.096 (*Arousal*) and 0.08 (*Valence*) which evidences that SVM was better than Naïve Bayes in terms of *accuracy*. The biggest difference between the two methods was observed in low partition of *Arousal* (55.3%) where *accuracy* was equal to 78% and 22.7% for SVM and Naïve Bayes correspondingly. The same test was applied for analyzing the performance in terms of *recall* and *true positive rate*. In this regard, no clear difference was observed between SVM and Naïve Bayes for *Arousal* (p-value = 0.307) and *Valence* (p-value = 0.324). This is due to the fact that Naïve Bayes had a superior performance in low partitions (biggest difference = 19.1%) whilst SVM was evidently better in medium (biggest difference = 71.1%) and high (biggest difference = 60.1%) ranges. Regarding the comparison in terms of *false positive rate*, no clear discrepancy was seen between the classification methods for both *Arousal* (p-value = 0.473) and *Valence* (p-value = 0.526). This is because Naïve Bayes had a lower *false positive rate* in medium partitions (biggest difference = 9.0%) whilst SVM performed better in low (biggest difference = 67.1%) and high (biggest difference = 7.8%) ranges.

The results are more interesting in terms of the bipartition scheme for SVM (refer to Table 6). The paired sample t test derived from the results of Table 6 evidenced that the percentage of correctly classified instances in SVM was statistically higher than that offered by Naïve Bayes in both *Arousal* (p-value = 0.196) and *Valence* (p-value = 0.239). The most significant gap between these algorithms can be found in low range of *Arousal* (54.6%) where *accuracy* was equal to 92.4% and 37.8% for SVM and Naïve Bayes respectively. The same analysis was implemented for verifying the *recall* and *true positive rate* of both algorithms under a bipartition labeling scheme. In this respect, no significant difference was observed between SVM and Naïve Bayes (p-value = 0.256). This is underpinned by the fact that Naïve Bayes had a superior performance in low partition of *Valence* (difference = 16.5%) while SVM was evidently better in *Arousal* (biggest difference = 61.4%) and high partition of *Valence* (difference = 75.1%) ranges. When analysing *false positive rate*, no clear discrepancy was seen between the classification methods (p-value = 0.256). Such finding is explained by tha fact that Naïve Bayes had a lower *false positive rate* in the high partition of *Valence* (difference = 16.5%) whilst SVM performed better in *Arousal* (biggest difference = 61.4%) and the low partition of *Valence* (difference = 75.1%).

On the other hand, the average accuracy using the bipartition labeling scheme was proved to be significantly higher than that provided using the tripartition labeling scheme for both *Arousal* (p-value = 0.014) and *Valence* (p-value = 0.003). When classifying *Arousal*, the best result using the bipartition scheme was 96.5% (high partition) whilst the best *accuracy* value using the tripartition scheme was 80.7% (medium partition). Similarly, upon considering *Valence* the best value in bipartition scheme was obtained in low partition (91.8%) which is higher than that achieved from the tripartition method scheme (81.5%). Average *recall* and *true positive rate* using the bipartition scheme were also concluded to be greater than those resulting from the use of tripartition scheme for *Arousal* (p-value = 0.04) and *Valence* (p-value = 0.024). When considering *Arousal*, the best values provided by the use of bipartition and tripartition schemes were 96.6% (high partition) and 86% (low partition) respectively. With respect to *Valence*, the highest score was obtained using the bipartition scheme (95.0%), which is greater than the best value obtained using the tripartition scheme (86.9%). Another aspect to be considered in this analysis is the *false positive rate*. In this regard, the t test evidenced that there is no statistically significant difference between the partitioning methods in both *Arousal* (p-value = 0.064) and *Valence* (p-value = 0.169) variables AQ6 (Fig. 4).

Figures 5, 6 illustrate the Receiver Operating Characteristic (ROC) curves for Arousal and Valence when using Naïve Bayes with statistical and powerband parameters. ROCs related to SVM are presented in Figs. 7, 8. When

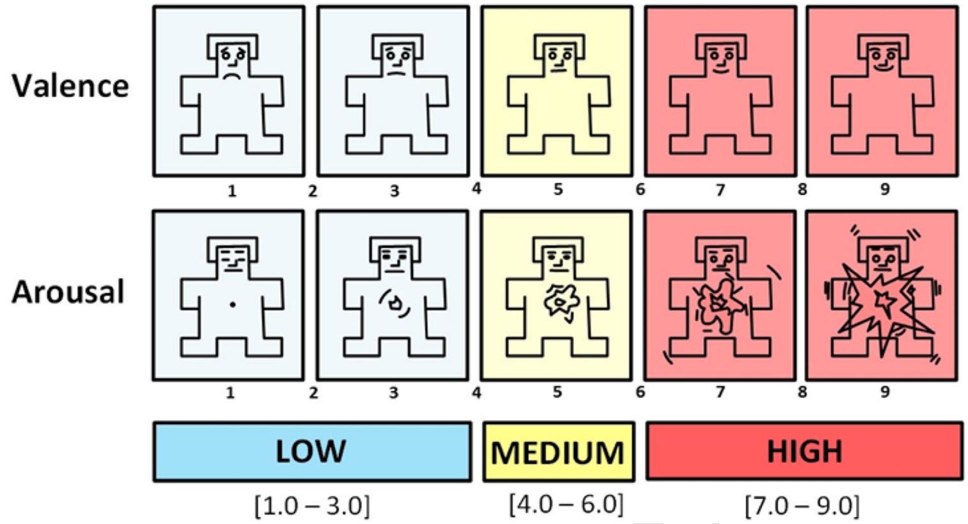


Fig. 4 Mapping from SAM scale Valence and Arousal values to Labels (Low, Medium, High) (Menezes et al. 2017)

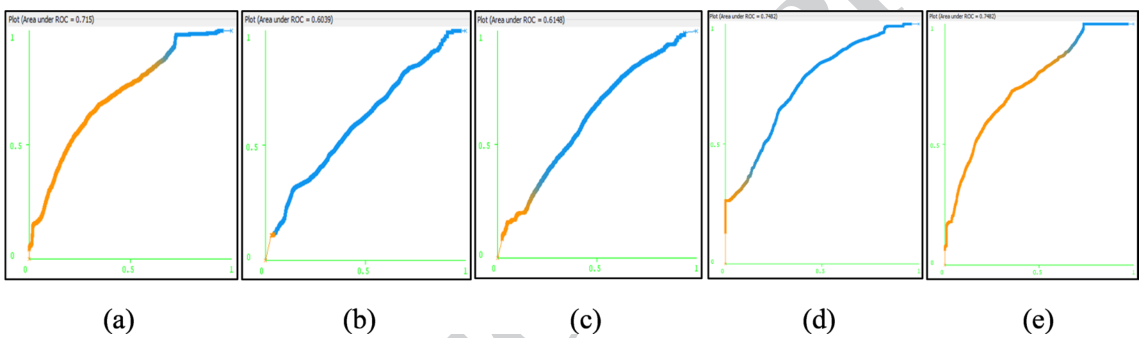


Fig. 5 ROC curves using Naïve Bayes with statistical and powerband parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)

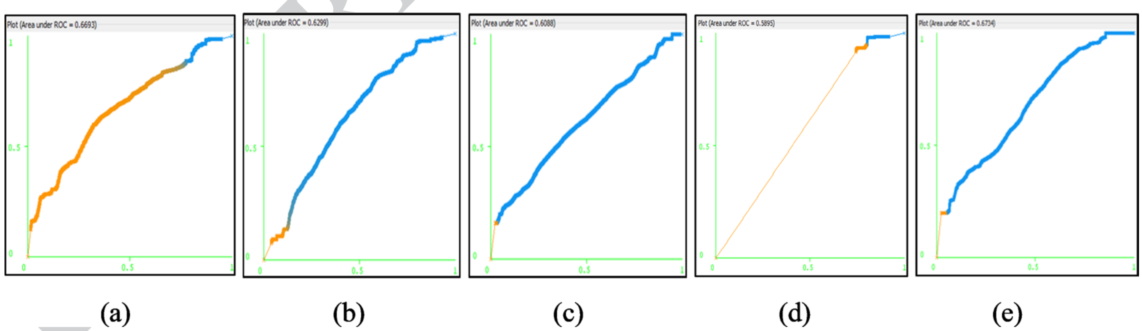


Fig. 6 ROC curves using Naïve Bayes with statistical and powerband parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

analyzing these curves, it can be corroborated that, in this case, SVM performs better than Naïve Bayes regarding Arousal. For instance, the area under curve in low partition (tripartition labeling scheme) of Arousal when using Naïve Bayes (0.715) (refer to Fig. 5a) is lower compared to SVM (0.8772) (refer to Fig. 7a). Similarly, when applying the bipartition labeling scheme and Naïve Bayes (refer to Fig. 5e), the area under curve in high partition of Arousal was 0.7482; however, when employing SVM, the area was found to be 0.9333 (refer to Fig. 7e). A similar conclusion was achieved when comparing the ROC curves in terms of Valence. For example, the area under curve in

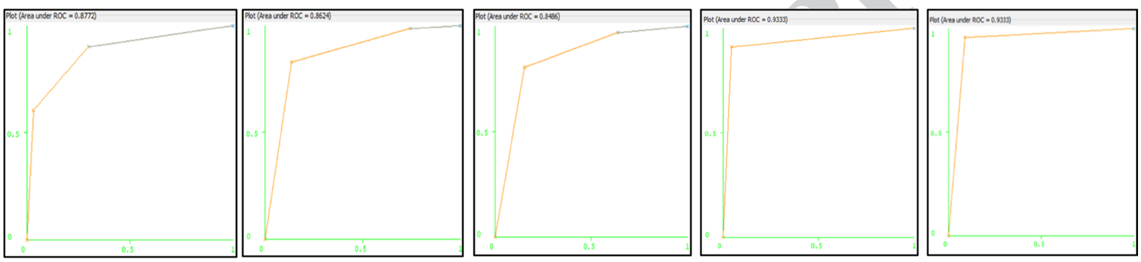
Journal : Large 12652	Article No : 1065	Pages : 20	MS Code : AIHC-D-18-00347	Dispatch : 26-9-2018

medium level (refer to Fig. 8b) when employing SVM was 0.8484, whilst Naïve Bayes provided an inferior performance (0.6299) (refer to Fig. 6b). In bipartition scheme, the area under ROC for the high partition was 0.6734 in Naïve Bayes (refer to Fig. 6e) and 0.8891 in SVM (refer to Fig. 8e).

SVM and Naïve Bayes were also tested by considering the powerband parameters derived from the EEG signals and employing the two partitioning schemes (refer to Tables 7, 8). In accordance with the resulting p-values for *Arousal* (p-value = 0.652) and *Valence* (p-value = 0.634), there is no significant difference between the classification algorithms regarding *accuracy*. The same conclusion was reached for *recall* and *true positive rate* in both *Arousal* (p-value = 0.811) and *Valence* (p-value = 0.985) variables. Similarly, no discrepancy was found between SVM and Naïve Bayes regarding *false positive rate* (p-value-Arousal = 0.473; p-value-Valence = 0.982).

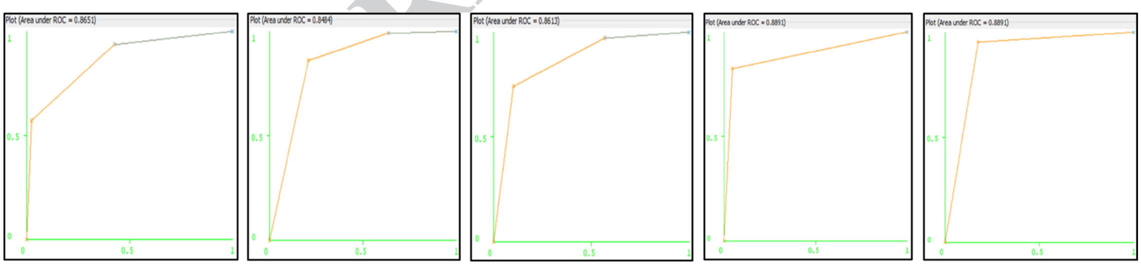
The bipartition labeling scheme was also implemented with powerband variables (refer to Table 8). The paired sample t test demonstrated that there are no meaningful differences when comparing *accuracy* values of SVM and Naïve Bayes (p-value [Arousal] = 0.486; p-value [Valence] = 0.945). Likewise, non-significant disparities were observed in *Arousal* (p-value = 0.821) and *Valence* (p-value = 0.980) when contrasting the algorithms in relation to *recall* and *true positive rate*. The same conclusion was obtained when correlating *false positive rates* (p-value [Arousal] = 0.821; p-value [Valence] = 0.980).

The average accuracy from the bipartition scheme was found to be statistically equivalent to that provided from the tripartition scheme regarding *Arousal* (p-value = 0.109). In



(a) (b) (c) (d) (e)

Fig. 7 ROC curves using SVM with statistical and powerband parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)



(a) (b) (c) (d) (e)

Fig. 8 ROC curves using SVM with statistical and powerband parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

Table 8 Results of classification process using bipartition labeling scheme (powerband parameters)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	0	0	0	0	66.8	17.8	17.8	6.6
	High	69.2	100	100	100	60.3	93.4	93.4	82.2
Naïve Bayes	Low	44.8	10.5	10.5	5.8	44.4	96.6	96.6	90.5
	High	70.2	94.2	94.2	89.5	79.1	9.5	9.5	3.4

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018

contrast, it was proved to be significantly higher in relation to *Valence* values (p-value = 0.006). When classifying emotion along *Valence* dimension, the best accuracy obtained using the bipartition scheme was 79.1% (high partition). Meanwhile, the best *accuracy* rate obtained using the tripartition scheme was 67.4% (high partition). Differences respecting average *recall* and *true positive rate* using the bipartition scheme were also investigated and confirmed to be non-significant in comparison with those emanating from the use of tripartition scheme for *Arousal* (p-value = 0.169) and *Valence* (p-value = 0.121). We also examined the false positive rates of both partitioning schemes. In this respect, p-values were determined to be greater than the alpha level and therefore, they do not present a meaningful statistical difference p-value [Arousal] = 0.187) and p-value [Valence] = 0.107) parameters.

Figures 9 and 10 present the ROC curves for Arousal when applying Naïve Bayes and SVM with powerband parameters respectively. ROCs related to Valence dimension are shown in Figs. 11 and 12. These plots evidence that, in most of these cases, Naïve Bayes provides better results than SVM in terms of Arousal. For example, the area under curve in low partition (tripartition scheme) of Arousal was 0.6164 when implementing Naïve Bayes

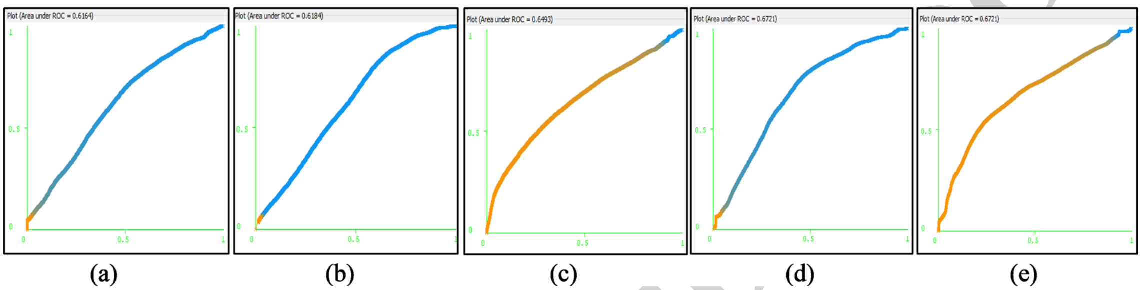


Fig. 9 ROC curves using Naïve Bayes with powerband parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)

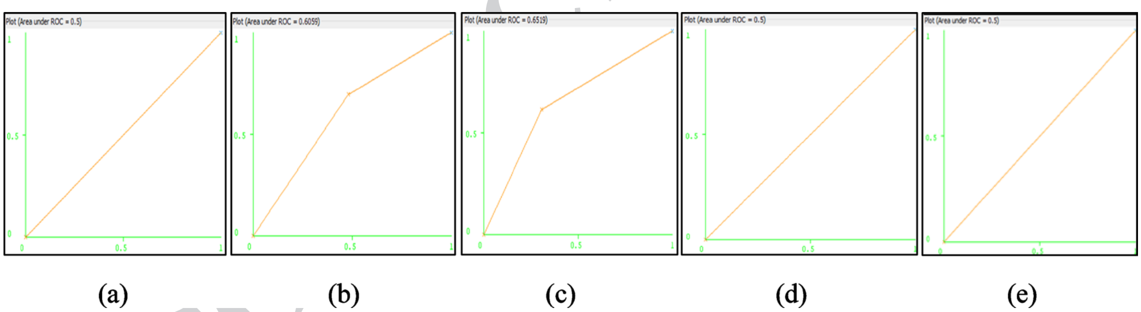


Fig. 10 ROC curves using SVM with powerband parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)

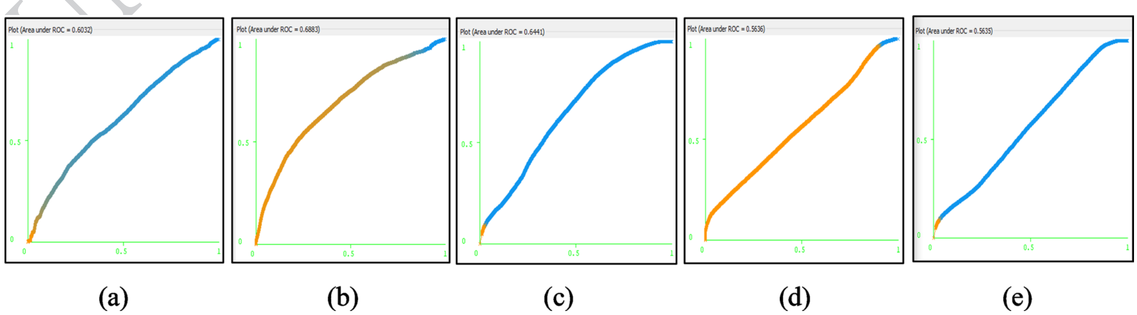


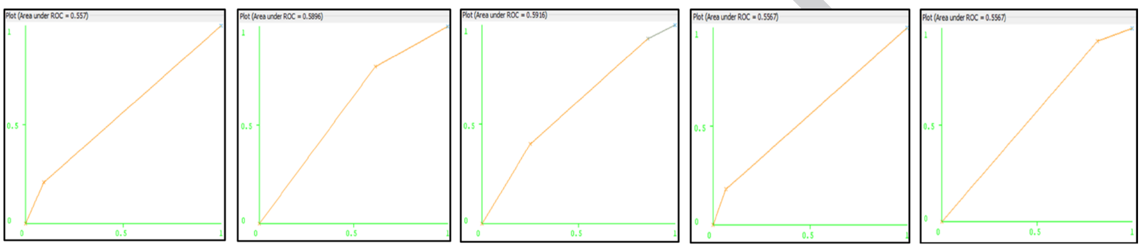
Fig. 11 ROC curves using Naïve Bayes with powerband parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

Journal : **Large 12652**	Article No : **1065**	Pages : **20**	MS Code : **AIHC-D-18-00347**	Dispatch : **26-9-2018**

(refer to Fig. 9a) and 0.5 when applying SVM (refer to Fig. 10a). Likewise, when using bipartition and Naïve Bayes (refer to Fig. 9d), the area under curve in low partition of Arousal was 0.6721; meanwhile, when employing SVM, the area was estimated to be 0.5 (refer to Fig. 10d). The only case where a different conclusion was drawn (SVM was better than Naïve Bayes) can be observed in the high level of tripartition (refer to Figs. 9c, 10c). On the other hand, when contrasting the classification methods in terms of Valence, it was also evidenced that Naïve Bayes was superior to SVM. In tripartition, for instance, the area under ROC in medium level when employing Naïve Bayes (refer to Fig. 11b) was 0.6883, while the performance provided by SVM was 0.5916 (refer to Fig. 12b). In bipartition, a small difference in favor of Naïve Bayes (0.0132) was observed between the areas under curve for the high partition: Naïve Bayes (refer to Fig. 11e) and SVM (refer to Fig. 12e).

The classification algorithms were also investigated and compared when using all the statistical features that were previously established in Sect. 4.3.1. Both the bipartition (refer to Table 9) and tripartition (refer to Tables 10, 11, 12) labeling schemes were also implemented. The p-values for *Arousal* (p-value = 0.182) and *Valence* (p-value = 0.416) show that there is no meaningful differences between the methods with respect to *the percentage of correctly classified instances*. The same conclusion was achieved for *recall* and *true positive rate* in both *Arousal* (p-value = 0.739) and *Valence* (p-value = 0.771) dimensions. Likewise, no dissimilarities were observed between SVM and Naïve Bayes in relation to *false positive rate* (p-value-Arousal = 0.477; p-value-Valence = 0.566).

The bipartition approach was also employed with the data derived from the predefined statistical parameters (refer to Table 10). Comparisons were also made using paired t tests. There were no differences in the mean *accuracy*



(a) (b) (c) (d) (e)

Fig. 12 ROC curves using SVM with powerband parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

Table 9 Results of classification process using tripartition labeling scheme (all statistical features)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	74.6	66.3	66.3	5	75.8	62.6	62.6	5.9
	Medium	81.0	83.2	83.2	13.5	78.4	86.1	86.1	20.8
	High	78.9	80.6	80.6	15	75.8	74.3	74.3	10.4
Naïve Bayes	Low	223	88.1	88.1	68.4	25.5	89	89	76.4
	Medium	62.2	11.3	11.3	4.7	52.7	15.4	15.4	12.2
	High	50.4	25.4	25.4	17.3	69.1	16	16	3.1

Table 10 Results of classification process using bipartition labeling scheme (all statistical features)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	90.2	91.1	91.1	4.4	89.5	81.7	81.7	7.1
	High	96.0	95.6	95.6	8.9	87.2	92.9	92.9	18.3
Naïve Bayes	Low	37.4	91.3	91.3	68.0	46.9	95.2	95.2	80.5
	High	89.2	32.0	32.0	8.7	84.6	19.5	19.5	4.8

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018

scores obtained by using SVM and Naïve Bayes (p-value [Arousal] = 0.418; p-value [Valence] = 0.461). Also, no critical discrepancies were seen in *Arousal* (p-value = 0.502) and *Valence* (p-value = 0.616) upon contrasting the methods regarding *recall* and *true positive rate* measures. This inference was further reached when comparing *false positive rates* (p-value [Arousal] = 0.502; p-value [Valence] = 0.616).

The mean *accuracy* from use of the bipartition scheme was concluded to be statistically bigger than that offered from the tripartition scheme regarding *Arousal* (p-value = 0.016) and *Valence* values (p-value = 0.003). When classifying affective *Arousal* dimension, the best *accuracy* score using the bipartition scheme was 96% (high partition) whilst the best value using the tripartition scheme was 81% (high partition). On the other hand, when categorizing *Valence*, the higher *percentage of correctly classified instances* using the bipartition scheme was 95.2% while use of the tripartion scheme provided 78.4%. However, when analyzing the differences between the bipartition and tripartition schemes in terms of average *recall* and *true positive rate*, no clear difference was detected in both *Arousal* (p-value = 0.082) and *Valence* (p-value = 0.062). We also investigated the false positive rates of the partitioning methods under study. The p-values were confirmed to be higher than 0.05 and hence, a meaningful statistical difference can not be underpinned (p-value [Arousal] = 0.150; p-value [Valence] = 0.093).

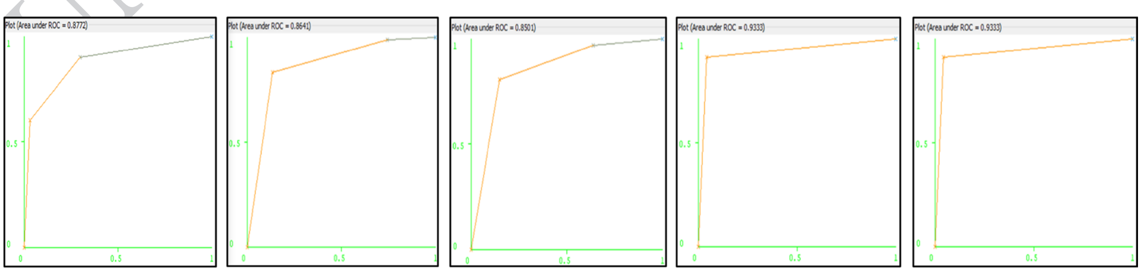
Figures 13 and 14 show the ROC plots for Arousal when implementing SVM and Naïve Bayes with statistical parameters correspondingly. The performance curves related to Valence parameter are presented in Figs. 15 and 16. These graphs demonstrate that, for this particular case, SVM performs better than Naïve Bayes in terms of Arousal. In particular, the area under curve in medium level (tripartition scheme) of Arousal was 0.8641 upon

Table 11 Results of classification process by using tripartition labeling scheme (significant statistical features)

Method	Level	Arousal				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	90.5	76.8	76.8	4.2	82.1	69.9	69.9	5.2
	Medium	98.3	96.3	96.3	11.4	84.9%	96.2	96.2	17.5
	High	95.7	93.3	93.3	12.6	82.1	83.6	83.6	9.1
Naïve Bayes	Low	34.0	94.8	94.8	63.2	34.8	83.1	83.1	71.0
	Medium	94.8	12.1	12.1	4.3	71.9	14.3	14.3	11.3
	High	76.8	27.3	27.3	16.0	94.3	14.9	14.9	2.9

Table 12 Results of classification process bipartition labeling scheme (significant statistical features)s

Method	Level	Arousalss				Valence			
		Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)	Accuracy (%)	Recall (%)	Tp rate (%)	Fp rate (%)
Support vector machine	Low	91.9	92.6	92.6	4.3	84.2	77.7	77.7	6.7
	High	97.8	97.1	97.1	8.7	82.0	88.4	88.4	17.4
Naïve Bayes	Low	39.7	94.8	94.8	65.4	46.1	83.1	83.1	68.7
	High	94.8	33.2	33.2	8.4	83.1	17.0	17.0	4.1



(a) (b) (c) (d) (e)

Fig. 13 ROC curves using SVM with statistical parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)

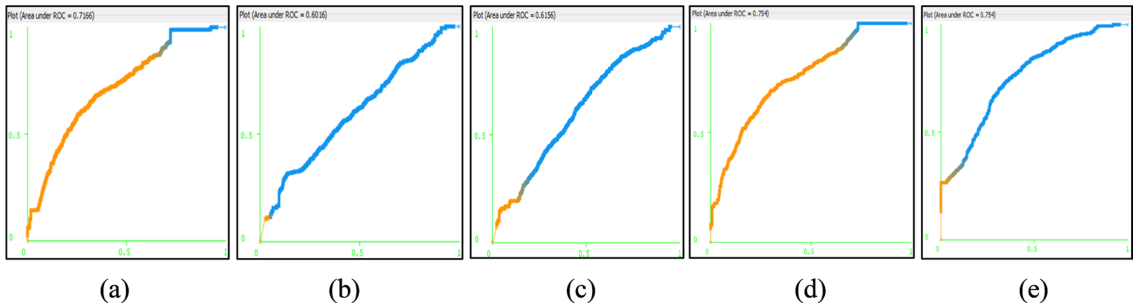


Fig. 14 ROC curves using Naïve Bayes with statistical parameters for **a** low, **b** medium, **c** high partitions of Arousal (tripartition labeling scheme) and **d** low, **e** high levels of Arousal (bipartition labeling scheme)

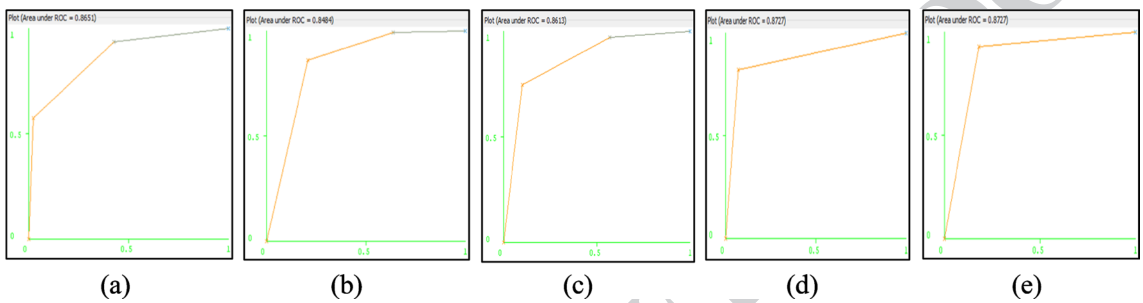


Fig. 15 ROC curves using SVM with statistical parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

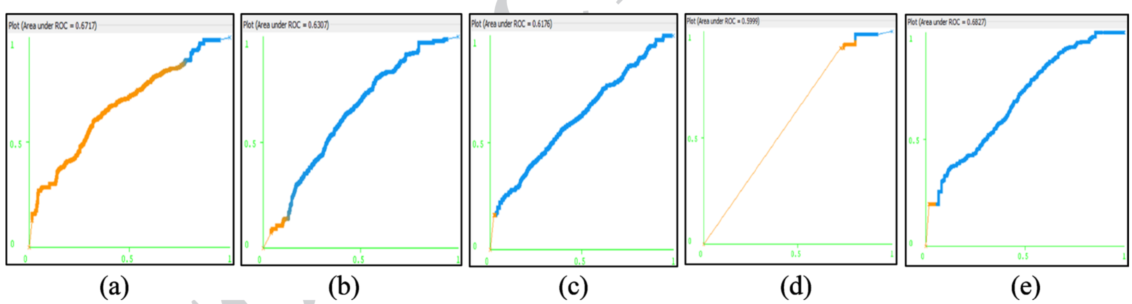


Fig. 16 ROC curves using Naïve Bayes with statistical parameters for **a** low, **b** medium, **c** high partitions of Valence (tripartition labeling scheme) and **d** low, **e** high levels of Valence (bipartition labeling scheme)

utilizing SVM (refer to Fig. 13b) and 0.6016 when executing Naïve Bayes (refer to Fig. 14b). In addition, when applying the bipartition approach and SVM (refer to Fig. 13d), the area under curve in low level of Arousal was 0.9333; however, when using Naïve Bayes, the area was calculated to be 0.754 (refer to Fig. 14d). When correlating SVM and Naïve Bayes in terms of Valence, it was also proved that SVM was better than Naïve Bayes. In tripartition, for example, the area under ROC curve in low level when applying SVM (refer to Fig. 15a) was 0.8651; nevertheless, the achieved performance in Naïve Bayes was 0.6717 (refer to Fig. 16a). Similarly, when using bipartition scheme, SVM (refer to Fig. 15d, e) performed better than Naïve Bayes and SVM (refer to Fig. 16d, e) for both low and high partitions of Valence.

Table 10 (tripartition) and 11 (bipartition) illustrate the results of classification metrics for both *Arousal* and *Valence* when using only significant statistical features. Compared to the results derived from the use of all the predefined statistical parameters, it was proved that the average *accuracy* in *Arousal* can be significantly increased when introducing only $\mu_{x_{\beta(AF3)}}$, $\mu_{x_{\theta(AF3)}}$ and $k_{x_{\beta(f3)}}$ (p-value = 0.002). Furthermore, it was found that the *recall* and *true positive rate* can be also augmented with the inclusion of the above-mentioned features (p-value = 0.004). On the other hand, a p-value = 0.007

Journal : **Large 12652**	Article No : **1065**	Pages : **20**	MS Code : **AIHC-D-18-00347**	Dispatch : **26-9-2018**

evidenced that a reduced false positive rate can be achieved with this change. In contrast, upon considering *Valence* dimension, there were no meaningful differences regarding accuracy when including $\sigma_{x_{\delta(AF4)}}$ and $k_{x_{\delta(AF3)}}$ (p-value = 0.092). In addition, the same conclusion was reached when analyzing the *recall/true positive rate* (p-value = 0.848) and false positive ratio (p-value = 0.052).

When contrasting the results emanating from significant statistical features and those resulting from powerband parameters, it was proved that significant parameters provided better accuracy of the *Arousal* (p-value = 0.002) and *Valence* (p-value = 0.001). The comparison in terms of *recall* and *true positive rate* was also studied. The results (p-value [Arousal] = 0.172; p-value [Valence] = 0.110) demonstrated that there is no clear difference between the scores derived from the aforementioned variables. A similar conclusion was drawn when comparing *false positive ratios* (p-value [Arousal] = 0.337; p-value [Valence] = 0.121).

Finally, it was found that the *percentage of correctly classified instances* was higher for *Arousal* when considering significant statistical parameters compared to that obtained upon combining powerband parameters and all statistical features (p-value = 0.002); although, no significant difference was found regarding *Valence* (p-value = 0.122). This relation was also examined by analyzing the *recall* and *true positive rate* which was concluded to be bigger in *Arousal* (p-value = 0.018) when using the significant parameters whilst no difference was detected in *Valence* (p-value = 0.585). The *false positive rates* did not differ significantly (p-value [Arousal] = 0.055; p-value [Valence] = 0.087).

Also, the parameters that can be better linked with the *Arousal* dimension are $\mu_{x_{\beta(AF3)}}$ (p-value = 0.041), $\mu_{x_{\theta(AF3)}}$ (p-value = 0.003) and $k_{x_{\beta(f3)}}$ (p-value = 0.029) whilst in *Valence*, the best features were $\sigma_{x_{\delta(AF4)}}$ (p-value = 0.034) and $k_{x_{\delta(AF3)}}$ (p-value = 0.048). In each case, combining significant variables improves the classification performance metrics. In this particular case, the results have revealed that fear, sadness and disgust were more difficult to discriminate. In this regard, other statistical and powerband features can be considered in order to increase the ability of distinguishing these emotions. Additionally, other brain positions may be better correlated to these emotional states and should be then further explored. In contrast, happiness, surprise and anger were found to be easier for detection.

6 Conclusions

Affective recognition is an important research area because it has potential to contribute to multiple applications in medicine, education and other fields. In accordance with the reported literature, several authors have applied DM, machine learning and artificial intelligence techniques for affective recognition (e.g. Support Vector Machine and Bayesian Networks).

Most previous works have made use of benchmark datasets where EEG signals are collected under controlled conditions that are very different from activities of everyday life. This study shows that satisfactory results can be observed when using EEG signals for affective recognition using a small headset with only 4 four channels and during activities that are typical of everyday life.

The results herein described can be potentially used for recognizing affective states. Considering significant statistical features combined with a bipartition labeling scheme, emotions can be effectively distinguished. Results show that SVM performed better than Naïve Bayes in some cases. Particularly, the highest percentage of correctly classified instances was achieved when using significant statistical parameters ([Arousal] = 98.3%; [Valence] = 94.3%). Additionally, the best recall/true positive rate ([Arousal] = 97.1%; [Valence] = 96.2%) and the lowest false positive ratio ([Arousal] = 4.2%; [Valence] = 2.9%) were also reached with the above-mentioned parameters.Furthermore, the bipartition approach was proved to be better than tripartition.

The above-mentioned results validate the ability of the SVM method for affective recognition when integrating with DM techniques. Another important aspect is that the use of statistical features plays a relevant role to increase the power and effectiveness of the proposed approach. In this regard, it was possible to provide an evidence base on the association between the significant features and emotional states which was concluded to be highly correlated with 94.90% and 85.08% for *Valence* and *Arousal* correspondingly, in addition to demonstrating their high predictive ability (94.86% and 83.10% respectively). Likewise, kurtosis was concluded to be highly correlated with both *Valence* and *Arousal* and it should be then used in future related studies.

Another relevant aspect is that most of the significant statistical parameters are related to β (*Arousal*) and δ (*Valence*) frequency bands. Furthermore, it was found that *AF*3 was identified as the most contributing position for affective recognition.

These results are extensible to medicine and education fields but also open to further questions that we aim to investigate. For example, could we use the most contributing electrode, AF3, and still have results that are interesting for context-aware application? How can we compare these results with the other signals in the dataset? Can we use the results obtained with the EEG data as a groudtruth for AQ8 analyzing other biosignals? Do the images obtained during the data collection match the results from the EEG and the Self Assessment Manikin? Or can we obtain with the biosignals a more accurate affective state evaluation other than the emotion that the person is willing to share with their facial expressions?

Acknowledgements The Authors which to acknowledge support from the REMIND Project from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 734355. The authors would also like to thank COST for supporting the work presented in this paper (COST-STSM-TD1405- 33385) and CNPq for the Science Without Borders Scholarship.

References

Abadi MK, Subramanian R, Kia SM, Avesani P, Patras I, Sebe N (2015) DECAF: MEG-based multimodal database for decoding affective physiological responses. IEEE Trans Affect Comput 6(3):209–222

Atkinson J, Campos D (2016) Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. Expert Syst Appl 47:35–41

Aymerich-Franch L (2010) Presence and emotions in playing a group game in a virtual environment: the influence of body participation. Cyberpsychol Behav Soc Netw 13(6):649–654

Barakat N, Bradley AP (2010) Rule extraction from support vector machines: a review. Neurocomputing 74(1–3):178–190

Barrett LF, Lewis M, Haviland-Jones JM (eds) (2016) Handbook of emotions. Guilford Publications, New York

Barrios MO, Jiménez HF (2015) Reduction of average lead time in outpatient service of obstetrics through six sigma methodology. In: Ambient intelligence for health. Springer, Cham, pp 293–302

Bastos-Filho TF, Ferreira A, Atencio AC, Arjunan S, Kumar D (2012) Evaluation of feature extraction techniques in emotional state recognition. In Intelligent human computer interaction (IHCI), 2012 4th international conference on IEEE, pp 1–6

Bekele E, Wade J, Bian D, Fan J, Swanson A, Warren Z, Sarkar N (2016) Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). In: Virtual reality (VR), 2016 IEEE, pp 121–130

Binder JR, Conant LL, Humphries CJ, Fernandino L, Simons SB, Aguilar M, Desai RH (2016) Toward a brain-based componential semantic representation. Cogn Neuropsychol 33(3–4):130–174

Botella C, Quero S, Baños RM, Perpiña C, Garcia-Palacios A, Riva G (2004) Virtual reality and psychotherapy. Cybertherapy 99:37–52

Boulos MNK, Hetherington L, Wheeler S (2007) Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. Health Inform Libr J 24(4):233–245

Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Therapy Exp Psychiatry 25:49–59

Brahnam S, Jain LC (2011) Virtual Reality in Psychotherapy, Rehabilitation, and Neurological Assessment. In: Advanced computational intelligence paradigms in healthcare 6. Virtual reality in psychotherapy, rehabilitation, and assessment. Springer, Berlin, pp 1–9

Chai TY, Woo SS, Rizon M, Tan CS (2010) Classification of human emotions from EEG signals using statistical features and neural network. In: International, vol. 1, No. 3. Penerbit UTHM, pp 1–6

Chatchinarat A, Wong KW, Fung CC (2017) Rule extraction from electroencephalogram signals using support vector machine. In: Knowledge and Smart Technology (KST), 2017 9th International Conference on IEEE, pp 106–110

De la Hoz E, de la Hoz E, Ortiz A, Ortega J, Martínez-Álvarez A (2014) Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps. Knowl-Based Syst 71:322–338

Desmet P (2018) Measuring emotion: development and application of an instrument to measure emotional responses to products. In: Funology 2. Springer, Cham, pp 391–404

Gerber AJ, Posner J, Gorman D, Colibazzi T, Yu S, Wang Z, … Peterson BS (2008) An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46(8):2129–2139

Glantz K, Rizzo AS, Graap K (2003) Virtual reality for psychotherapy: current reality and future possibilities. Psychotherapy 40(1–2):55

Hodges LF, Anderson P, Burdea GC, Hoffmann HG, Rothbaum BO (2001) Treating psychological and phsyical disorders with VR. IEEE Comput Graphics Appl 21(6):25–33

Ip HHS, Byrne J, Cheng SH, Kwok RCW (2011) The samal model for affective learning: a multidimensional model incorporating the body, mind and emotion in learning. In: 17th International Conference on Distributed Multimedia Systems, DMS 2011. Knowledge Systems Institute Graduate School

Islam M, Ahmed T, Mostafa SS, Yusuf MSU, Ahmad M (2013) Human emotion recognition using frequency & statistical measures of EEG signal. In: Informatics, Electronics & Vision (ICIEV), 2013 International Conference on IEEE, pp 1–6

Izquierdo-Reyes J, Ramirez-Mendoza RA, Bustamante-Bello MR, Navarro-Tuch S, Avila-Vazquez R (2017) Advanced driver monitoring for assistance system (ADMAS). Int J Interact Design Manufac (IJIDeM), 1–11

Jenke R, Peer A, Buss M (2014) Feature extraction and selection for emotion recognition from EEG. IEEE Trans Affect Comput 5(3):327–339

Jerritta S, Murugappan M, Nagarajan R, Wan K (2011) Physiological signals based human emotion recognition: a review. In: Signal processing and its applications (CSPA), 2011 IEEE 7th International Colloquium on IEEE, pp 410–415

Jirayucharoensak S, Pan-Ngum S, Israsena P (2014) EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. Sci World J

Katsigiannis S, Ramzan N (2017) Dreamer: a database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices. IEEE J Biomed Health Inform

Kim J, André E (2008) Emotion recognition based on physiological changes in music listening. IEEE Trans Pattern Anal Mach Intell 30(12):2067–2083

Kim YE, Schmidt EM, Migneco R, Morton BG, Richardson P, Scott J, et al (2010) Music emotion recognition: a state of the art review. In: Proc. ISMIR, pp 255–266

Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, et al (2012) Deap: a database for emotion analysis; using physiological signals. IEEE Trans Affect Comput 3(1):18–31

Konstantinidis EI, Frantzidis CA, Pappas C, Bamidis PD (2012) Real time emotion aware applications: a case study employing emotion evocative pictures and neuro-physiological sensing enhanced by Graphic Processor Units. Comput Methods Progr Biomed 107(1):16–27

Lan Z, Sourina O, Wang L, Liu Y (2014) Stability of features in real-time EEG-based emotion recognition algorithm. In: Cyberworlds (CW), 2014 International Conference on IEEE, pp 137–144

Lan Z, Sourina O, Wang L, Liu Y (2016) Real-time EEG-based emotion monitoring using stable features. Vis Comput 32(3):347–358

Lin YP, Wang CH, Jung TP, Wu TL, Jeng SK, Duann JR, Chen JH (2010) EEG-based emotion recognition in music listening. IEEE Trans Biomed Eng 57(7):1798–1806

Liu Y, Sourina O (2013) EEG databases for emotion recognition. In: Cyberworlds (CW), 2013 International Conference on IEEE, pp 302–309

Liu Y, Sourina O (2014) Real-time subject-dependent EEG-based emotion recognition algorithm. In: Transactions on computational science XXIII. Springer, Berlin Heidelberg, pp 199–223

Liu Y, Sourina O, Nguyen MK (2010) Real-time EEG-based human emotion recognition and visualization. In: Cyberworlds (CW), 2010 International Conference on IEEE, pp 262–269

Maaoui C, Pruski A (2010) Emotion recognition through physiological signals for human-machine communication. In: Cutting edge robotics, InTech

Mampusti ET, Ng JS, Quinto JJI, Teng GL, Suarez MTC, Trogo RS (2011) Measuring academic affective states of students via brainwave signals. In: Knowledge and systems engineering (KSE), 2011 Third International Conference on IEEE, pp 226–231

Menezes MLR, Samara A, Galway L, Sant'Anna A, Verikas A, Alonso-Fernandez F et al (2017) Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset. Pers Ubiquit Comput 21(6):1003–1013

Murugappan M, Rizon M, Nagarajan R, Yaacob S, Hazry D, Zunaidi I (2008a) Time-frequency analysis of EEG signals for human emotion detection. In: 4th Kuala Lumpur international conference on biomedical engineering, Springer, Berlin, pp 262–265

Murugappan M, Rizon M, Nagarajan R, Yaacob S, Zunaidi I, Hazry D (2008b) Lifting scheme for human emotion recognition using EEG. In: Information Technology, 2008. ITSim 2008. International Symposium on IEEE, vol. 2, pp 1–7

Murugappan M, Juhari MRBM, Nagarajan R, Yaacob S (2009) An Investigation on visual and audiovisual stimulus based emotion recognition using EEG. Int J Med Eng Inform 1(3):342–356

Murugappan M, Ramachandran N, Sazali Y (2010) Classification of human emotion from EEG using discrete wavelet transform. J Biomed Sci Eng 3(04):390

National Research Council (1999) How people learn: Brain, mind, experience and school, Committee on Developments in the Science of Learning. In: Bransford, JD, Brown, AL, Cocking, RR

Niedermeyer E, da Silva FL (eds) (2005) Electroencephalography: basic principles, clinical applications, and related fields. Lippincott Williams & Wilkins

Nugent C, Synnott J, Gabrielli C, Zhang S, Espinilla M, Calzada A et al (2016) Improving the quality of user generated data sets for activity recognition. In: Ubiquitous computing and ambient intelligence. Springer, Cham, pp 104–110

Ontiveros-Hernández NJ, Pérez-Ramírez M, Hernández Y (2013) Virtual reality and affective computing for improving learning. Res Comput Sci 65:121–131

Parsons TD, Rizzo AA (2008) Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J Behav Therapy Exp Psychiatry 39(3):250–261

Petrantonakis PC, Hadjileontiadis LJ (2010) Emotion recognition from EEG using higher order crossings. IEEE Trans Inf Technol Biomed 14(2):186–197

Picard RW, Vyzas E, Healey J (2001) Toward machine emotional intelligence: analysis of affective physiological state. IEEE Trans Pattern Anal Mach Intell 23(10):1175–1191

Picard RW, Papert S, Bender W, Blumberg B, Breazeal C, Cavallo D et al (2004) Affective learning—a manifesto. BT Technol J 22(4):253–269

Pool E, Brosch T, Delplanque S, Sander D (2016) Attentional bias for positive emotional stimuli: a meta-analytic investigation. Psychol Bull 142(1):79

Riva G, Mantovani F, Capideville CS, Preziosa A, Morganti F, Villani D et al (2007) Affective interactions using virtual reality: the link between presence and emotions. Cyber Psychol Behav 10(1):45–56

Rizon M, Murugappan M, Nagarajan R, Yaacob S (2008) Asymmetric ratio and FCM based salient channel selection for human emotion detection using EEG. WSEAS Trans Signal Process 4(10):596–603

Schaaff K, Schultz T (2009) Towards an EEG-based emotion recognizer for humanoid robots. In: Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on IEEE, pp 792–796

Schlögl A, Slater M, Pfurtscheller G (2002) Presence research and EEG. In: Proceedings of the 5th International Workshop on Presence, vol. 1, pp 9–11

Shu Y, Wang S (2017) Emotion recognition through integrating EEG and peripheral signals. In: Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on IEEE, pp 2871–2875

Soleymani M, Lichtenauer J, Pun T, Pantic M (2012) A multimodal database for affect recognition and implicit tagging. IEEE Trans Affect Comput 3(1):42–55

Sreeshakthy M, Preethi J (2016) Classification of Human Emotion from Deap EEG Signal Using Hybrid Improved Neural Networks with Cuckoo Search. BRAIN Broad Res Artif Intell Neurosci 6(3–4):60–73

Subasi A (2007) EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst Appl 32(4):1084–1093

Vijayan AE, Sen D, Sudheer AP (2015) EEG-based emotion recognition using statistical measures and auto-regressive modeling. In: Computational intelligence & communication technology (CICT), 2015 IEEE International Conference on IEEE, pp 587–591

Wang Q, Sourina O (2013) Real-time mental arithmetic task recognition from EEG signals. IEEE Trans Neural Syst Rehabil Eng 21(2):225–232

Wang XW, Nie D, Lu BL (2011) EEG-based emotion recognition using frequency domain features and support vector machines. In: International Conference on Neural Information Processing. Springer, Berlin, pp 734–743

Wu D, Courtney CG, Lance BJ, Narayanan SS, Dawson ME, Oie KS, Parsons TD (2010) Optimal arousal identification and classification for affective computing using physiological signals: virtual reality stroop task. IEEE Trans Affect Comput 1(2):109–118

Wu S, Wang S, Zhu Y, Gao Z, Yue L, Ji Q (2016) Employing subjects' information as privileged information for emotion recognition from EEG signals. In: Pattern recognition (ICPR), 2016 23rd International Conference on IEEE, pp 301–306

Zhong Y, Jianhua Z (2017) Subject-generic EEG feature selection for emotion classification via transfer recursive feature elimination. In: Control Conference (CCC), 2017 36th Chinese IEEE, pp 11005–11010

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal : Large 12652 | Article No : 1065 | Pages : 20 | MS Code : AIHC-D-18-00347 | Dispatch : 26-9-2018